



Porcine adenoviruses types 1, 2 and 3 have short and simple early E-3 regions¹

P. Seshidhar Reddy, Neeraja Idamakanti, J. Brian Derbyshire, Éva Nagy*

Department of Veterinary Microbiology and Immunology, University of Guelph, Guelph, Ontario N1G 2W1, Canada

Received 11 August 1995; accepted 4 January 1996

Abstract

The nucleotide sequence of the E-3 region genes, the hexon associated protein pVIII genes, and part of the fiber genes coding for the N-terminal tail regions, of porcine adenovirus (PAV) types 1 and 2 were determined. The sizes of the E-3 regions were found to be 1162 and 1222 bp, respectively. The five open reading frames (ORF) encoded within the sequenced regions of PAV types 1 and 2 shared a high degree of homology with the published sequences of the corresponding ORFs of PAV-3. The E-3 regions of PAV types 1, 2 and 3 were further characterized by Northern blot analysis and 5' and 3' end mapping of the transcripts by S1 nuclease analysis. The results of these experiments indicated that the E-3 regions in these three viruses are shorter and simpler in organization than the E-3 regions of human adenoviruses. A potential promoter for the E-3 regions of these PAVs was identified.

Keywords: Porcine adenovirus; Shorter and simpler early region 3

1. Introduction

Porcine adenoviruses (PAVs) are currently classified into five serotypes (Derbyshire et al., 1975; Hirahara et al., 1990). The proposed use of PAVs as viral vaccine vectors, particularly for the

induction of mucosal immune responses against enteric viral antigens in pigs, has stimulated interest in the molecular genetics of these viruses. The cloning and physical mapping of the DNA of serotypes 1-5 (Reddy et al., 1993, 1995b; Kleiboeker et al., 1993; Tuboly et al., 1995) and the nucleotide sequence analysis of the inverted terminal repeats (ITR) of all five serotypes (Reddy et al., 1995c), and of the E-3 region of PAV-3 and PAV-4 (Reddy et al., 1995a; Kleiboeker, 1994) have been reported.

* Corresponding author.

¹ The nucleotide sequence information presented in this paper is deposited with the GenBank database under the accession numbers L-43364, L-43365 and U10433.

While PAV types 1-3 are readily distinguished by virus neutralization tests, with no significant cross reactivity (Derbyshire et al., 1975), all were isolated from the same region of the U.K. during the same period (Clarke et al., 1967), and the data available on their restriction enzyme analysis and ITR sequences indicate a close genetic relationship (Garves and Xuan, 1989; Reddy et al., 1995c).

We are proposing to use PAV 1-3 as expression vectors for foreign genes in swine. The construction of recombinant PAV-1, 2 and 3, each expressing a foreign antigen gene from a pathogen would enable the use of one serotype for primary immunization and the other two serotypes for subsequent immunizations. As such the immunity against one PAV serotype would not interfere with the replication of another serotype expressing the same foreign gene. The sequence analysis of the E-3 region and the neighbouring pVIII and fiber regions of the PAV-3 genome was described recently (Reddy et al., 1995a). In the present paper we report the nucleotide sequence analysis of the pVIII, E-3, and part of the fiber coding regions of the genomes of PAV-1 and PAV-2, and Northern blot analysis and 5' and 3' end mapping of the E-3 transcriptional unit of the E-3 regions of PAV 1-3.

2. Materials and methods

2.1. Viruses and preparation of viral DNA

The 25R strain of PAV-1, the A47 strain of PAV-2 and the 6618 strain of PAV-3 were propagated in secondary pig kidney (PK) cells. Intracellular viral DNA was extracted by the method of Hirt (1967).

2.2. Primers, PCR and DNA sequencing

The sequence of the pVIII and fiber genomic regions of PAV-3 (Reddy et al., 1995a) was used to design primers specific to the 3' end of the pVIII and the 5' end of the fiber genes. These were synthesized at the Regional DNA Synthesis Laboratory, University of Calgary, Calgary. The

sequence of the primer specific to the 3' end of pVIII was 5'-ACGGACTCTGTAGCAGG-3', and that of the 5' end of the fiber gene was 5'-GTAGACTGGATCGAAG-3'. The E-3 regions of PAV-1-3 were amplified by the polymerase chain reaction (PCR), using approximately 10 ng of DNA template, 30 pmol of each of the two primers, and 2.5 U *Taq* polymerase (BRL) with the following temperatures: denaturation at 94°C (1 min), annealing at 55°C (1 min) and extension at 72°C (1 min). The products were analyzed on agarose gels, purified, blunt ended with T4 DNA polymerase, and cloned into pGEM-7Zf (+). A nested population of overlapping deletion clones was made using exonuclease III and the sequence of one strand was determined by the dideoxy nucleotide chain termination technique using these deletion clones. The sequence determination of the other strand and of both strands of the gene coding for pVIII and part of the fiber coding region was done from the cloned genomic fragments at the Mobix Sequence Facility, Hamilton, Ontario.

2.3. RNA preparation

PK cells grown in Petri dishes were infected with 5-10 plaque forming units (pfu) per cell of PAV-1, 2 or 3 and incubated for 2-16 h. Total cellular RNA was extracted from mock infected or PAV infected cells with acid guanidinium thiocyanate-phenol-chloroform mixture as described by Chomczynski and Sacchi (1987). Polyadenylated RNA was selected by passing RNA over oligothymidylic acid cellulose columns.

2.4. Northern analysis

The total RNA extracted from the mock and virus infected cells was separated on 1% agarose-formaldehyde gels and transferred to Nytran membranes (Sambrook et al., 1989). The blots were prehybridized, hybridized and washed as described (Sambrook et al., 1989). The DNA fragments spanning the E-3 regions, labelled with [α -³²P]dCTP by the random primer labelling method, were used as probes.

2.5. Nuclease protection

To prepare the probes, the selected fragments of DNA were digested with the appropriate enzymes. They were either 5' end labelled with [γ -³²P]ATP, using polynucleotide kinase, or 3' end labelled with [α -³²P]dCTP with the Klenow fragment. The labelled DNA was digested with a second enzyme to remove the label at one end and the desired fragment was isolated from the gel. S1 nuclease protection assays were performed according to Berk and Sharp (1977). Approximately 200 μ g of RNA from mock or PAV infected cells 8 or 9 h post-infection (p.i.), depending on the serotype, were used. Hybridization temperatures were chosen depending on the GC content of the probe. Samples were treated with 200 U of S1 nuclease at 37°C for 1–2 h to digest the unhybridized portions of the RNA and DNA. S1 nuclease reactions were analyzed on 6% acrylamide/8 M urea sequencing gels.

3. Results

3.1. Features of the established sequences

Two primers were used to amplify the E-3 regions of PAV-1 and PAV-2 and the products were approximately 1.1 kb each (data not shown). The polymerase chain reaction (PCR) products were cloned into plasmids. Sequencing of these, and the appropriate cloned genomic fragments of putative pVIII and fiber genes, viz. the *Kpn*I–*Sna*BI and *Sna*BI–*Bam*HI fragments of PAV-1 and the *Kpn*I–*Bam*HI and *Bam*HI D fragments of PAV-2, revealed that the sizes of the E-3 regions of PAV-1 and PAV-2 were 1162 and 1222 bp, respectively. The DNA sequences of the r-strands are given in Fig. 1. The sequenced region from PAV-1 had five open reading frames (ORF) on the r-strand with a coding potential for more than 50 amino acids (aa), while PAV-2 also had five ORFs which could encode polypeptides of 44 or more amino acids. The number of ORFs on the l-strand varied from two in PAV-1 to four in PAV-2. A TATA box was located between nucleotides (nt) 356 and 359 within the body of ORF-1

in both PAV-1 and PAV-2. One canonical polyadenylation signal was found at the end of the E-3 of PAV-1 and PAV-2 (Fig. 2). The nucleotide sequences of PAV-1 and PAV-2 were compared with each other and to the corresponding published sequence of PAV-3 (Reddy et al., 1995a). The nucleotide sequence homology between the sequenced regions of PAV-1 and PAV-3 was 96.1%, while between PAV-1 and PAV-2, and between PAV-2 and PAV-3, the homologies were 91.7% and 91.3%, respectively.

ORFs 1–4 on the r-strand of PAV-1 had theoretical coding capacities for proteins of 24.4, 13.7, 27.1 and 9.7 kDa and those of PAV-2 for 24.4, 14.0, 28.5 and 10.5 kDa proteins respectively. ORFs 1–4 were overlapping, whereas ORF-5 was not, in either PAV-1 or PAV-2 (Fig. 2).

The deduced amino acid sequences of ORF-1 of PAV-1 and PAV-2 were compared with each other and to the sequence (Reddy et al., 1995a) of the PAV-3 ORF-1 (Fig. 3(a)). An amino acid identity of 98.6% between PAV-1 and PAV-3 and 95.0% between PAV-2 and PAV-3 was observed. The amino acid identity between the corresponding sequences of PAV-1 and PAV-2 was 96.4%.

The predicted amino acid sequence of the protein encoded by ORF-2 of PAV-1 and PAV-2 revealed an identity of 92.3% (Fig. 3(b)). These ORFs also showed amino acid identities of 97.4% and 94%, respectively, with the corresponding ORF of PAV-3. One potential glycosylation signal was present at the carboxyl terminus of the putative polypeptide in both serotypes. The deduced amino acid sequences of the proteins encoded by ORF-3 of PAV-1 and PAV-2 showed a high degree of identity (Fig. 3(c)), 81.6% between PAV-1 and PAV-2, and 70.2% between PAV-1 and PAV-3. The predicted amino acid sequence of ORF-3 had two membrane associated helices. In the case of the deduced amino acid sequence of ORF-4, the amino acid identity between PAV-1 and PAV-2 was 77.6% and between PAV-1 and PAV-3 the amino acid identity was 79.7% (Fig. 3(d)). For both ORF-3 and ORF-4 of PAV-1 and PAV-2, the putative ATGs conform to the Kozak rule for translation.

The total coding capacity of ORF-5 is unknown since only part of the sequence was deter-

PAV-1	ATGAGCAAAACAAATCTCCACCCCGTACATGGTGGTCTATCAGGCCAACTCTGGGCGTGCCTCGCCCGCTCGATTACTCCACCGCATGAAATTGGC	100
PAV-2	ATGAGCAAAACAAATCTCCACCCCGTACATGGTGGTCTATCAGGCCAACTCTGGGCGTGCCTCGCCCGCTCGATTACTCCACCGCATGAAATTGGC	100
	DVIII start codon	
PAV-1	TCAAGTGCCTGGGCTTCCATGATTGGTCAGGTCAATGACATCGGACACACCCAGGAACAAATTCTCATTCGGCCAGGCCCTTATCACCGAGACGCCACGCC	200
PAV-2	TCAAGTGCCTGGGCTTCCATGATTGGCGGTCAATGCCATCGGACAAAGACAGGAACCAAGATTCTCATTCGGCCAGGCCCTTATCACCGAGACGCCACGCC	200
PAV-1	CGTCCAAATCCCCATCTGGGCCCGCCGCCCTCTGGCTCAGATGACGGAAACCGCCACCCACCTGCACTCTGGGCCGCCACAGAGATTGGAGGCCAGA	300
PAV-2	CGTCCAAATCCCCATCTGGGCCCGCCGCCCTCTGGCTCAGATGACGGAAACCGCCACCCACCTGCACTCTGGGCCGCCACAGAGATTGGAGGCCAGA	300
PAV-1	CTGACTGACGCCGGCATGCAATTAGCGGGGGCGGGGCCCTGGCCCGACAGACTTATATGGCCCTGACCCCTCGCGGCCAGAGGATTCAGCTAACGGGG	400
PAV-2	CTGACTGACGCCGGCATGCAATTAGCGGGGGCGGGGCCCTGGCGGCCAGAGACTTATATGGCCCTGACCCCTCGCGGCCAGAGGATTCAGCTAACGGGG	400
	GC box CAT box GC box TATA box T. start site	
PAV-1	ACCTGCGCTCTCTGGGGAGCACTCTCCGCCGAGGCCATCTTCCAGCTCGGGAGGGCTTGCGGCCCTCGCTCCTCACACCCACCGACGCCACTCGACCGT	500
PAV-2	ACCTGCGCTCTCTGGGGAGCACTCTCCGCCGAGGCCATCTTCCAGCTCGGGAGGGCTTGCGGCCCTCGCTCCTCACACCCACCGACGCCACTCGACCGT	500
PAV-1	GCAGAACTCCAGCTCCCTGCCCGAGGGCGCATCGGAGCAACCTTGTGGCGAGTCTGCTACCCCGTCTACATCAACCCCTCTCCGGACCG	600
PAV-2	GCAGAACTCCAGCTCCCTGCCCGAGGGCGCATCGGAGCAACCTTGTGGCGAGTCTGCTACCCCGTCTACATCAACCCCTCTCCGGACCG	600
PAV-1	CCCCGGACATACCCGACCGATTCATCCCAACTACAAACATCCTCACCGACTCTGTAGCAGGCTATGACTGACGGTCTCCCAAGGGCCACGACGTG-----	695
PAV-2	CCCCGGACATACCCGACCGATTCATCCCAACTACAAACATCCTCACCGACTCTGTAGCAGGCTATGACTGACGGTCTCCCAAGGGCCACGACGTG-----	700
	ORF 3 start codon DVIII stop codon	
PAV-1	-TGGCGGGAGCTCTCGACCAGCACCGCGGCCAGTGCCTAACCGCTGTGCT-----TGGCGGGAGGGATCACCGGAGTACTTTTGCACTACCCCGCG	791
PAV-2	CTGGGGGACCTCTTGGACCGACGCCGGCCAGTGCTTCTAACCGCTGTGCTTGCGGCCAGGATGGGATCGCCGGGGAGTACTTTTGCACTACCCCGCG	800
PAV-1	AGCACCTTGAGGCCGAGTGCAATTCCCGACTCTCTCGCAAGAGGCCACGGPCGCGCTTCAGGCTCCCCAGCGCTCACACGACGCCGGCACCGCATGG	891
PAV-2	AGCACCTTGAGGCCGAGTGCAATTCCCGACTCTCTCGCAAGAGGCCACGGPCGCGCTTCAGGCTCCCCAGCGCTCACAGGGACGCCGGCACCGCATGG	900
	ORF 3 start codon	
PAV-1	AGAACGCCACCATCTCTCACTTCGTAACACTTGCGGCCCTCTTCAAGGTCCGCTGCTCTCGGCCATCTCTGCTCTCACCCCTCTCTCTCGACCTT	991
PAV-2	AGAACGCCACCATCTCTCACTTCGTAACACTTGCGGCCCTCTTCAAGGTCCGCTGCTCTCGGCCATCTCTGCTCTCACCCCTCTCTCTCGACCTT	1000
PAV-1	CTGTGTGAGCTGTACAACCCGCTGTAGCGTCAAGCCCTACACTTCCCTCGCGTCCAGTTCTGAAACGAGATAAGACCGGACTCCAGCTCGGCTCTGGC	1091
PAV-2	CTGTGTGAGCTGTACAACCCGCTGTAGCGTCAAGCCCTACACTTCCCTCGCGTCCAGTTCTGAAACGAGATAAGACCGGACTCCAGCTCGGCTCTGGC	1100
	ORF 2 stop codon	
PAV-1	TCTCGAGACCGATGATGTGAGATT-----TTGAGCTGACCAAGGACACACCCAAACGGAGACATCTAGGCGACATAGTCATCAACAAACCGATGGGGCCCA	1188
PAV-2	TCTGGGTTCTGGGATATGAGAGATGACATTACCTGTCTACCGAGAACCCGAGATGACCTCTAGGTTGATCTGTCATCAACAAACCGATGGGGCCCA	1200
PAV-1	AGACCTGGGCCCTGGGATACCTCTATGGCGTCCGGCAGTTCTCTCTCGCCATCATCATCATCGTCTCATCTCTACTACCGCCGCTACGTGCTGGC	1288
PAV-2	AGACCTGGGCCCTGGGATACCTCTATGGCGTCCGGCAGTTCTCTCTCGCCATCATCATCATCTCATCTACTACCGCCGCTACGTGCTGGC	1300
	ORF 4 start codon	
PAV-1	CACCGGCCCTCATCTGGCAGGCCAGATGGTCTCCGGGCCCTCTCCGGAAACACTCTGGGCCACAGCTCATGGTTACTCCCTAAACAA-----	1382
PAV-2	TACCGGCCCTCATCTGGCAGGCCAGATGGTCTCCGGGCCCTCTCCGGAAACACTCTGGGCCACAGCTCATGGTTACTCCCTAAACAA-----	1397
PAV-1	GTCAACCCCTGCAACTGCTCTCGCCGCTTGGAGGAGATGGTCTCTACTACACAAACCTGGCTCTCATGGCCCTGGGGGCTCATCTCTGCTCACCG	1482
PAV-2	GTCAACCCCTGCAACTGCTCTCGCCGCTTGGAGGAGATGGTCTCTACTACACAAACCTGGCTCTCATGGCCCTGGGGGCTCATCTCTGCTCACCG	1497
PAV-1	CCATGGTCGCCCTAGGCCACTGGGATAGGGACCAAGATGCCCAAGCAAGGACCGC-----GCCCGAC-----CACCCCTCACCTATGTGGGGCCAGGCC	1570
PAV-2	CCATGGTCGCCCTGCCCAACTGGGATAGGGACCAAGATGCCCAAGCAAGGACCGC-----GCCCGAC-----CACCCCTCACCTATGTGGGGCCAGGCC	1597
	ORF 4 stop codon	
PAV-1	CGAGGACCATCTACGATGAGCCA-----ACGG-----TAGGGCAATACGATGAGCAACTGAGCTGGCTCCCT-----TTTCCCCCTTC-----CC	1646
PAV-2	CGAGGACCATCTACGATGAGCCCCATCACGGGGGGTCCCAAGGCTGGGAGCTACATGAGATGAAACTGAGACCTCCCGCTTC-----CCATCCCCCTCTCC	1697
	T. stop site ORF 3 stop codon	
PAV-1	CCATTCTCTATTTTC-----AATAAGTTGCTACCTGAGTTCATCCACACTCGGTCTGCCGTGCACTCTATCCATGCGGTTTTCCATACTCAC	1738
PAV-2	CCATTCTCTTGTGACATCTTAAAGTTACTACCTGACTCATCCACACTCGGTCTGCCGTGCACTCTATCCATGCGGTTTTCCATACTCAC	1797
	Poly A signal T. stop site	
PAV-1	ATAGCGCAGCGGCCACGCCCTGCCAGGTGAGGAACACTGAAAGTAACATTG-----GCTCTGTCACAGCAGCCCTGATAAACCACTGTCACCATG	1837
PAV-2	ATAGCGCAGCGGCCACGCCCTGCCAGGTGAGGAACACTGAAAGTAACATTG-----GCTCTGTCACAGCAGCCCTGATAAACCACTGTCACCATG	1897
	Fiber start codon	
PAV-1	GGACCGAGAGAGCAGAGCGCGAGCTCCCGAGGACTTCGATGCCAGCTCTACCCCTATGACGCCCGCAGCTGAGATCAATCCACCCCTCGTCAGGGGG	1937
PAV-2	ACGGCGAGAGAGCAGAGCGCGAGCTCCCGAGGACTTCGATGCCAGCTCTACCCCTATGACGCCCGCAGCTGAGATCAATCCACCCCTCGTCAGGGGG	1997
PAV-1	ACGGATTTCCACCAATCCGTGGACGGGTG 1966	
PAV-2	ACGGCTTCACCAATCCGTGGACGGGTG 2026	

Fig. 1. Alignment of nucleotide sequences of the pVIII, E-3 and part of the fiber genes of PAV-1 and PAV-2. The GC, CAAT and TATA boxes and AATAAA signals are underlined. Transcriptional start (T. start site) and stop (T. stop site) sites and translation initiation and termination sites are indicated.

BEST AVAILABLE COPY

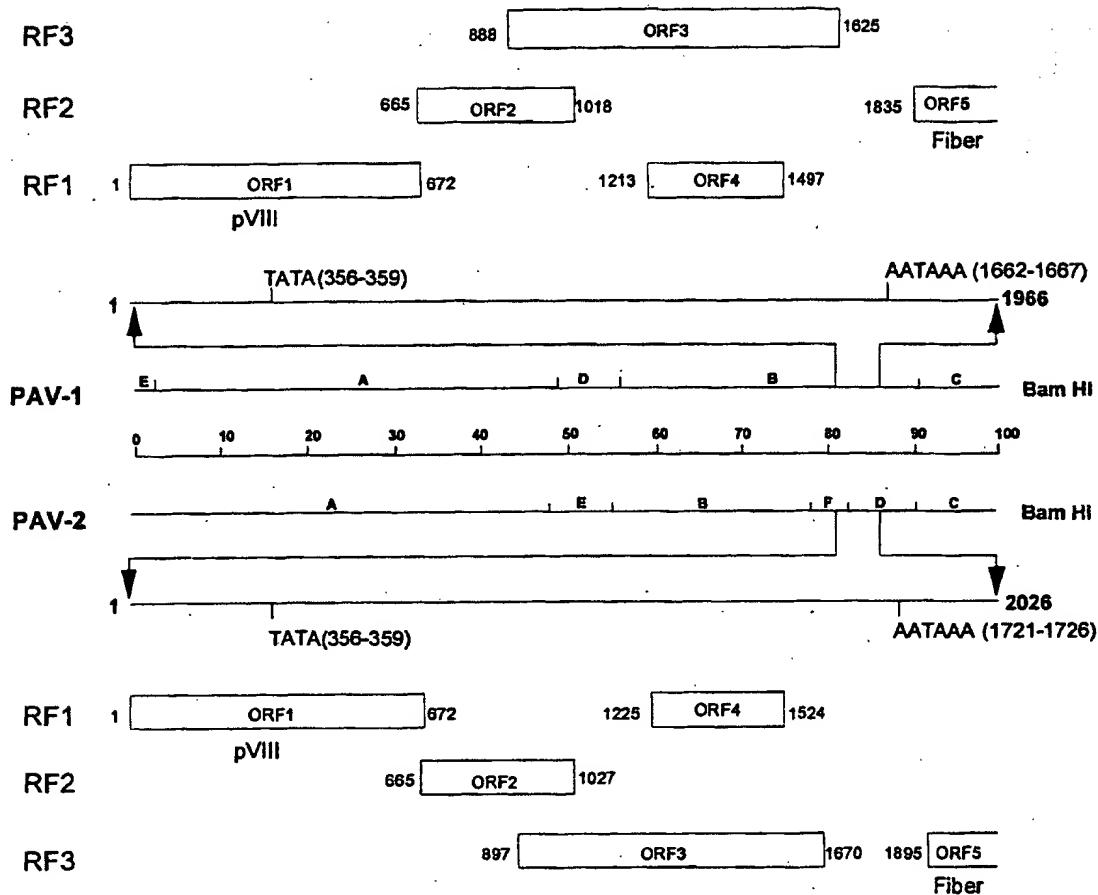


Fig. 2. Physical maps of the genomes of PAV-1 and PAV-2 for *Bam*HI. The sequenced regions of the genomes are enlarged and the ORFs on the r-strand with a coding capacity for 44 or more amino acids, and the location of TATA and AATAAA signals, are indicated.

mined. In this region an identity of 90.7% was observed between the deduced amino acid sequences of ORF-5 of PAV-1 and PAV-2 (Fig. 3(e)). When compared with the amino terminal region of the fiber of PAV-3, amino acid identities of 95.3 and 90.7% were found between PAV-1 and PAV-3, and PAV-2 and PAV-3 respectively.

3.2. Northern blot analysis of E-3 mRNAs in PAV-1, 2 and 3

To determine the time of expression of the E-3 region genes, Northern blot analysis was carried out (Fig. 4). In all three serotypes only a single

band corresponding to approximately 1.6–1.9 kb was first detected by 6–8 h p.i. depending on the serotype. Transcription from this region continued up to at least 14 h p.i. Larger transcripts, 2.8 and 5.3 kb in size, were initially detected at 12 h p.i.

3.3. Nuclease protection assays

To map the 5' and 3' ends of the transcripts generated from the E-3 region to the DNA sequence, S1 nuclease protection assays were used. In the case of PAV-3, a probe (905 bp *Kpn*I–*Acc*I fragment) labelled at the 5' end was used to locate

(a)

PAV-1 ORF1 MSKQIPTPYMMWSYQPQSGRAAGASVVDYSTRMNWLSAGPSMIGQVNDIRHTRNQILIRQALITETPRFVQNPSSWPARFLPQMTOPPTHLHLPRNEILEGR 100
 PAV-2 ORF1 MSKQIPTPYMMWSYQPQSGRAAGASVVDYSTRMNWLSAGPSMIGQVNAIRQDNRQILIRQALITETPRFVQNPSSWPARFLPQMTOPPTHLHLPRNEILEGR 100
 PAV-3 ORF1 MSKQIPTPYMMWSYQPQSGRAAGASVVDYSTRMNWLSAGPSMIGQVNDIRHTRNQILIRQALITETPRFVQNPSSWPARFLPQMTOPPTHLHLPRNEILEGR 100

PAV-1 ORF1 LTDAGMQLAGGGALAPRDLYALTTLRGRGIOLNEDLPSASTLRPDGIFQLOGGGRSSSPNPTDAYLTQNSSSLPRSGGIGSDQFVRFVPTVYINPPSGP 200
 PAV-2 ORF1 LTDAGMQLAGGGALAPRDLYALTTLRGRGIOLNEDLPSASTLRPDGIFQLOGGGRSSSPNPTDAYLTQNSSSLPRSGGIGSDQFVRFVPTVYINPPSGP 200
 PAV-3 ORF1 LTDAGMQLAGGGALAPRDLYALTTLRGRGIOLNEDLPSASTLRPDGIFQLOGGGRSSSPNPTDAYLTQNSSSLPRSGGIGSDQFVRFVPTVYINPPSGP 200

PAV-1 ORF1 PGTYPDQFILANYNILTDHSVAGYD 223
 PAV-2 ORF1 PGTYPDQFILANYNILTDHSVAGYD 223
 PAV-3 ORF1 PGTYPDQFILANYNILTDHSVAGYD 223

(b)

PAV-1 ORF2 MTDGPQGQQL--LRELLDQHRRQCPNRCC-FAREGIHPEYFCITRSHPEACI PDSLQEGHGLRPSLPTRYNDRHRDGERTIL/TSYTCGPASFKVRCL 96
 PAV-2 ORF2 MTDGPQAOQQQQQLRDLDOHRRQCPNRCCPAREGIRPEYFCITRSHPEACI PDSLQEGHGLRPSLPTRYSDRERHDGERTIL/TSYTCGPASFKVRCL 99
 PAV-3 ORF2 MTDGPQGQQL--LRELLDQHRRQCPNRCC-FAREGIHPEYFCITRSHPEACI PDSLQEGHGLRPSLPTRYSDRERHDGERTIL/TSYTCGPASFKVRCL 96

PAV-1 ORF2 CGHPAPHPPLLDLICELYNRS 117
 PAV-2 ORF2 CSHPAPHPPLFDLICELYNRS 120
 PAV-3 ORF2 CGHPAPHPPLLDLICELYNRS 117

(c)

PAV-1 ORF3 METAPSSLRTTAARLLSRSAVSAAIILLLTLPPSTFCVSCITARSVSPYTSRVOFLNEIEPDSSSGSGSGIDDDDEDFE--LTSD-TPNEDILGSIVINN 96
 PAV-2 ORF3 METAPSSLRTTAARLLSKPAAACAAIILLLTEPSTFCVSCITARSVSPYTSRVOFLNEIEPDSSGSGSGIDDDDEDFE--YLSTETPPDDDLGGIVINN 97
 PAV-3 ORF3 METAPSSLRTTAARLLSKPAAVSAAIILLLTEPSTFCVSCITARSVSPYTSRVOFLNEIEPDSSGSGSGIDDDDEDFE--YLSTETPPDDDLGGIVINN 98

PAV-1 ORF3 QIGPKTLALGYFYAAQFVFFAI I I I I I I I I I I Y Y R Y V L A T A L I V O R Q M S S E A L R K T P S A T V M V T P P Q - - V T P C N C S C R F E D V V Y Y T T S V F M A L V G 193
 PAV-2 ORF3 QIGPKTLALGYFYAAQFVFLAI I I I I I I I I I I Y Y R Y V L A T A L I V O R Q M S S E A L R K T P A A T E V - P A R Q P A V T P C N C S C R F E D V V Y Y T T S G F M V L V G 195
 PAV-3 ORF3 QIGPKTLALGYFYAAQFVFFAI I I I I I I I I I I Y Y R Y V L A T A L I V O R Q M S S E A L R K T P S A T V V V T P P Q - - V T P C N C S C R F E D V V Y Y T T S V F M - - - 190

PAV-1 ORF3 ILLTAMVRLANIVDQMSRNRÄ--PPPLTIVGFSARDHLYDEPTVG-----QVYOMK 245
 PAV-2 ORF3 ILLTALVRLANIVDQMSRNRÄPPPLTIVGFSARDHLYDEPTVG-----QVYOMK 257
 PAV-3 ORF3 -----PWWASSSCSPPNSAMPT-----G 207

(d)

PAV-1 ORF4 MPPGSLSSSPSSSSSSSTTAATCWPSSCSARCPGPRPSGKPSRPOSWLLPONK-SPPATAPASRRNCSTTQPPSSWPNMSSCSP----- 90
 PAV-2 ORF4 MPPVMSWSPSSSSFSSTTAATCWPSSCSARCPGPRPSGKPSRPOSWLLPONK-SPPATAPASRRNCSTTQPPSSWPNMSSCSP----- 91
 PAV-3 ORF4 MPPCSLSSLPSSSSSSSTTAATCWPSSCSARCPGPRPSGKPSRPOSWLLPONK-SPPATAPASRRNCSTTQPPSSCPGGPHPPAHRHGPQQL 97

PAV-1 ORF4 -----PMS-----A 94
 PAV-2 ORF4 -----PMSASPTG 99
 PAV-3 ORF4 DSGSDAQEPRAAAATAPHLCOTLRRGPHL---R 128

(e)

PAV-1 ORF5 MGPKKQKRELPEPDPEVYFYDAPQLOINPPFVSGDGFHQSVDG 43
 PAV-2 ORF5 NTAKKQKRELPEPDPEVYFYDAPQLOINPPFVSGDGFHQSVDG 43
 PAV-3 ORF5 MGPKKQKRELPEPDPEVYFYDAPQLOINPPFVSGDGFHQSVDG 43

Fig. 3. Comparison among (a) the predicted amino acid sequences of ORFs 1 of PAV-1, PAV-2 and PAV-3; (b) the predicted amino acid sequences of ORF 2 of PAV-1, PAV-2 and PAV-3; (c) the predicted amino acid sequences of ORF 3 of PAV-1, PAV-2 and PAV-3; the potential membrane associated helices are underlined; (d) the predicted amino acid sequences of ORF 4 of PAV-1, PAV-2 and PAV-3; and (e) the predicted amino acid sequences of the amino terminal regions of the fiber proteins of PAV-1, PAV-2 and PAV-3, using the Sequence Alignment program: mismatch penalty 2, open gap penalty 4 and extended gap penalty 1. Potential glycosylation signals (N-X-S/T) are underlined.

the 5' end. The observed S1-resistant band (192–195 nucleotides (nt)) indicated that the initiation of transcription takes place 22–25 bp downstream from the 3' end of a TATA box located within the body of ORF-1 (Fig. 5, panel 3). In the case of PAV-1 and PAV-2, probes (243 bp *Apa*I–*Acc*I fragment) labelled at the 5' ends were used. In both serotypes two closely spaced protected fragments (50–51 nt) were observed (Fig. 5, panels 1 and 2), indicating that transcription in the E-3 region initiates 26–27 nt downstream from the

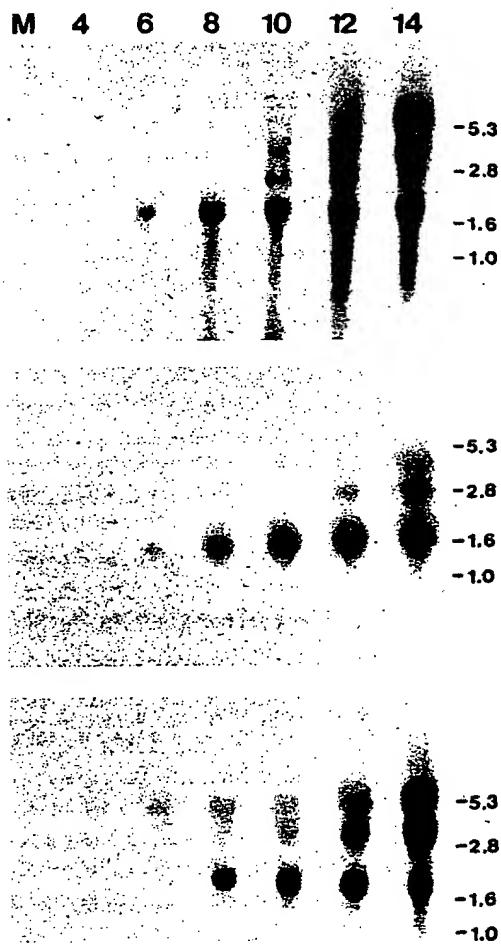


Fig. 4. Northern blot analysis of the E-3 region of PAV-1 (upper panel), PAV-2 (middle panel) and PAV-3 (lower panel). Total cellular RNA was isolated from mock (M) infected or virus infected cells (4–14 h p.i.). Positions of RNA molecular weight markers are indicated on the right with their size in kb.

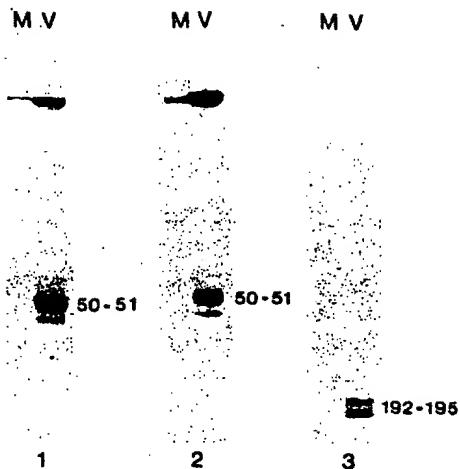


Fig. 5. 5' end mapping of mRNAs of the E-3 region of PAV-1, PAV-2 and PAV-3. For PAV-1 (1) and PAV-2 (2) the probes were 5' end labelled at the *Apa*I site of the *Apa*I–*Acc*I fragment. For PAV-3 (3) the *Kpn*I–*Acc*I fragment was labelled at the *Acc*I site. Lanes M and V represent the S1 nuclease protected fragments with RNA extracted from mock and PAV-infected cells, respectively. The sizes of the protected fragments (nt) are indicated on the right of each band.

TATA box at nt 385–386. Analysis of the sequence surrounding the transcriptional start sites showed a GC box, a CAAT box, and a GC box followed by a TATA box (Fig. 1).

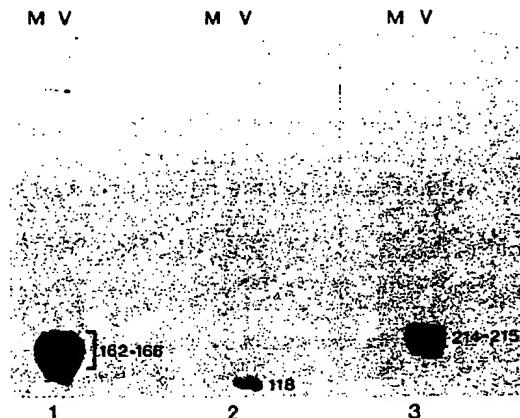


Fig. 6. 3' end mapping of PAV-1 (1), PAV-2 (2) and PAV-3 (3) E-3 mRNAs. The probe used is described in the text. Lanes M and V represent S1 nuclease protected fragments using RNA extracted from mock and PAV-infected cells, respectively. The sizes of the protected fragments (nt) are indicated on the right of each band.

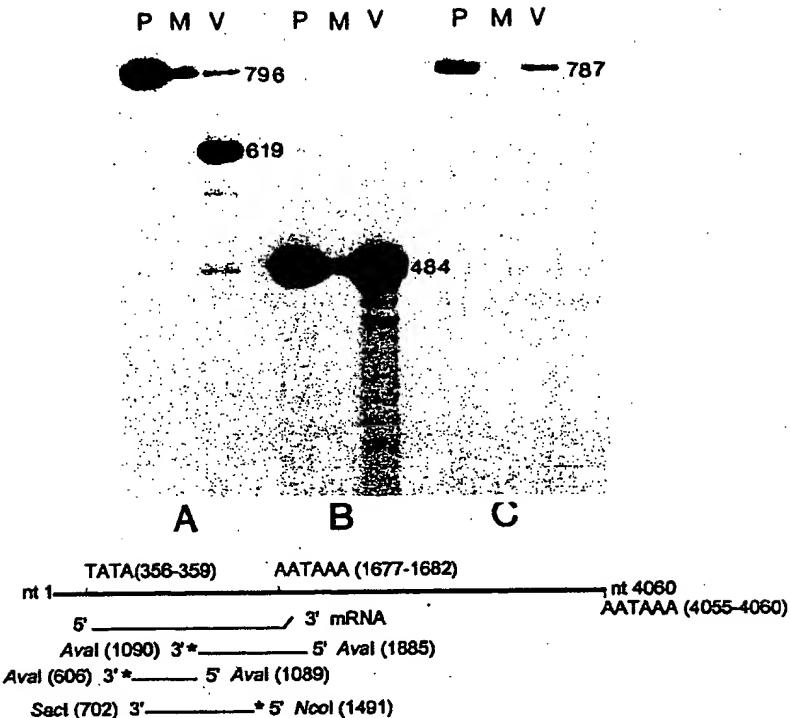


Fig. 7. Mapping of potential splice sites of the E-3 mRNAs of PAV-3. The probes used, shown diagrammatically below the mRNA line, were the 796 bp *Ava*I fragment (panel A) and the 484 bp *Ava*I fragment (panel B), labelled at the 3' end, and the 787 bp *Sac*I-*Nco*I fragment labelled at the 5' end (panel C). Lane P represents the probes alone, lanes M and V represent the S1 nuclease protected fragments with RNA extracted from mock and PAV-3 infected cells, respectively. The sizes of the protected fragments (nt) are indicated on the right of each band.

For 3' end mapping of transcripts of the E-3 region in PAV-3, the 3' end labelled 303 bp *Nco*I-*Taq*I fragment was used as a probe (Fig. 6, panel 3). The lengths of the protected fragments (214 and 215 nt) indicated that the transcripts generated from the E-3 region terminate 25–26 bp downstream from the 3' end of the E-3 polyadenylation signal at 1707–1708 nt. In the case of PAV-1 the 294 bp *Nco*I-*Taq*I fragment labelled at the 3' end was used as a probe and two closely spaced protected fragments (162 and 166 nt) indicated that transcription termination takes place 32–36 bp downstream from the 3' end of the AATAAA signal (nt 1699–1703). For 3' end mapping of the E-3 transcripts in PAV-2 the 3' end labelled 311 bp *Ava*II-*Acc*I

fragment was used as a probe and a 118 nt protected fragment indicated that transcriptional termination occurs 28 bp downstream from the 3' end of the polyadenylation signal (Fig. 6, panels 1 and 2).

To map the splice sites in transcripts of the E-3 region of PAV-3, a 796 bp *Ava*I fragment (nt 1090–1885) and a 484 bp *Ava*I fragment (nt 606–1089) both labelled at the 3' end, and a 787 bp *Sac*I-*Nco*I fragment (nt 701–1491), labelled at the 5' end, were used as probes (Fig. 7). The protected fragments consistently co-migrated with the original probes except with the *Ava*I 796 bp probe, which extended beyond the transcriptional stop site, in which the protected fragment migrated below the probe.

4. Discussion

The DNA sequence analysis showed a very high degree of homology (> 90%) across the sequenced E-3 regions among PAV-1, PAV-2 and PAV-3. It was reported earlier that the DNA homology among human adenoviruses within a subgroup was approximately 85–95%, with the exception of subgroups A and D, whose members share DNA homology of 70 and 20%, respectively (Sussenbach, 1984). Similar physical maps (Reddy et al., 1995b) and a high degree of DNA homology in the sequenced region among these three PAVs suggest that they belong to a single subgroup.

In human and animal adenoviruses, including PAV-3, but with the exception of ovine adenovirus, the E-3 region lies between the genes coding for the precursor protein pVIII and the fiber protein of the L-4 and L-5 regions of the genome, respectively (Cladaras and Wold, 1985; Reddy et al., 1995a; Vrati et al., 1995). In the case of PAV-1 and PAV-2, since the deduced amino acid sequence of ORFs 1 and 5 showed a high degree of identity with the putative pVIII and fiber proteins of PAV-3, it seems very likely that ORFs 1 and 5 are the genes coding for the pVIII and fiber proteins, respectively. Based on these observations the E-3 regions of PAV-1 (nt 673–1834) and PAV-2 (nt 673–1894) are 1162 and 1222 bp long, respectively. In PAV-3 and PAV-4 the lengths of the E-3 regions were found to be 1179 and 1879 bp, respectively (Reddy et al., 1995a; Kleiboecker, 1994). Relatively short E-3 regions (782 bp) have also been reported for mouse adenovirus type 1 (MAV-1) (Raviprakash et al., 1989), 1.1 kb for CAV-1 (Dragulev et al., 1991) and 1.5 kb for BAV-3 (Mittal et al., 1992). As reported earlier the E-3 regions of animal adenoviruses are variable in size, but are consistently smaller than the E-3 regions of HAV genomes (Cladaras and Wold, 1985), and our findings further confirm this characteristic of animal adenoviruses.

The present study demonstrated three overlapping ORFs in the E-3 regions of PAV-1 and PAV-2. The E-3 regions of PAV-3 (Reddy et al., 1995a), PAV-4 (Kleiboecker, 1994), and BAV-3

(Mittal et al., 1992) also had three overlapping ORFs. However, the E-3 region of HAV-2 and HAV-5 included ten ORFs (Cladaras et al., 1985). The similar sizes and almost equivalent positions of the three ORFs of the E-3 regions of PAV-1, PAV-2 and PAV-3 further suggest extensive similarity among the three serotypes. As in the E-3 regions of PAV-3 and PAV-4, the E-3 regions of PAV-1 and PAV-2 have only one polyadenylation signal. The E-3 regions of HAVs are divided into two 3' co-termination families depending on which polyadenylation signal is used (Cladaras et al., 1985). These features of the E-3 regions of human adenoviruses, in contrast to those of other animal adenoviruses, show that the E-3 regions of animal adenoviruses are much simpler in organization.

The putative pVIII genes in PAV-1, PAV-2 and PAV-3 were immediately followed by an ORF (ORF-2) with potential to encode a polypeptide of approximately 14 kDa. The nucleotide sequence at the junction between these two ORFs was ATGACTGA, thus the stop codon for the pVIII polypeptide and the start codon for the 14 kDa protein ORF overlap by 8 nt (Fig. 1). A sequence similar to ORF-2 of PAVs was also noticed in HAV-2 (Roberts et al., 1986), HAV-3 (Signas et al., 1986) and CAV-2 (Linne, 1992). Since this ORF is conserved in a number of adenoviruses, it is very likely to be functional. The deduced amino acid sequence of ORF-2 in PAV-1, PAV-2 and PAV-3 showed a high degree of identity (> 92%) among the three viruses, and a lesser degree of homology with the putative polypeptide encoded by ORF-2 of the E-3 region of CAV-2 (Linne, 1992). In CAV-2 and HAV-2 the ORF-2 of the E-3 region is followed by a putative L-4 polyadenylation signal (Cladaras et al., 1985; Linne, 1992) but no such signal was found in the E-3 regions of PAV-1, PAV-2 and PAV-3.

The predicted amino acid sequence of ORF-3 of the E-3 regions in PAV-1, PAV-2 and PAV-3 revealed two membrane associated helices. This 19 kDa E-3 protein of HAV-2, which binds to MHC class I antigens, also has two hydrophobic domains. A third apolar domain is located immediately C-terminal to the NCS consensus motif of PAV-1 and PAV-2.

The deduced amino acid sequence of ORF-4 of PAV-1 and PAV-2 showed a high degree of identity between the two serotypes and with the corresponding ORF of PAV-3. The functional significance of these ORFs is currently unknown. The N-terminus of the putative fiber protein, representing the tail region, was identified in ORF-5 by alignment of the deduced amino acid sequence with the N-terminal 43 amino acids of the PAV-3 fiber protein. The sequence PVYPYD, shown to be important for binding of the tail region of the fiber to the penton base (Devaux et al., 1987), was also detected in ORF-5 of PAV-1 and PAV-2 (Fig. 3(e)).

In HAV-2, transcription from the E-3 region was detected by 2 h p.i., and with the onset of DNA replication it increased 3-10 fold, which was attributed to an increase in the number of DNA templates (Shaw and Ziff, 1980). In the present study the Northern blot analysis indicated that transcription from the E-3 regions of PAV-1, PAV-2 and PAV-3 begins 6-8 h p.i. The levels of transcription, as in HAV-2 were found to increase up to 14 h p.i. (the last time period tested). The Northern blot studies indicated that transcription from the E-3 regions begins by 6 h p.i., which is prior to onset of PAV-3 DNA replication (data not shown). Consequently, as for other adenovirus E-3 regions, that for PAV was considered to be transcribed early. The 1.6-1.9 kb size estimated for the PAV-1 and PAV-2 E-3 transcripts was somewhat longer than the 1.3-1.4 kb size of the E-3 coding region. Since there is a polyadenylation site at the 3' ends of both E-3 genes and since all adenoviral transcripts are polyadenylated, the extra length noted for both E-3 transcripts could be due to the addition of a poly(A) tail.

In the 5' and 3' end mapping experiments, two or three closely spaced bands of protected DNA were seen. This could be either a consequence of S1 nuclease 'nibbling' into the RNA-DNA hybrids or it could be due to heterogeneous transcription initiation and termination sites. In nuclease experiments, multiple bands of nuclease protected fragments are usually seen because RNA-DNA hybrids breathe at the ends and nuclease 'nibbles' into the hybrid (Cladaras et al.,

1985). Heterogeneous transcription initiation sites were also reported for HAV-5, HAV-2 and other viruses (Cladaras et al., 1985).

In HAV-2, HAV-5 and MAV-1, transcriptional control elements were shown to be located upstream of the transcriptional initiation site of the E-3 region (Baker and Ziff, 1981; Cladaras et al., 1985; Beard et al., 1990). Lee et al. (1982) reported that the sequences upstream of the transcriptional start site were essential for efficient initiation of transcription of HAV-2 mRNA in vitro. The sequence upstream of the transcriptional initiation site in each of the three PAVs included a GC box, a CAAT box, and a GC box followed by a TATA box. A typical eukaryotic promoter also contains a similar structure. Therefore these sequence motifs might act as transcriptional control elements for the E-3 regions in PAVs. The polyadenylation signal AATAAA is required for the generation of 3' ends by cleavage of the primary transcript. The E-3 regions of PAV-1, PAV-2 and PAV-3 also had canonical polyadenylation signals upstream of the transcriptional termination sites.

The transcriptional start and stop sites of the E-3 regions of these PAVs indicated that the L-4 region overlaps the E-3 region, and the L-5 region starts just downstream of the E-3 region. A similar genomic organization was reported for human and murine adenoviruses (Cladaras et al., 1985; Beard et al., 1990).

In human and murine adenoviruses alternative splicing is used to generate overlapping mRNAs (Signas et al., 1986; Beard et al., 1990; Cladaras et al., 1985). In the case of PAV-3, S1 nuclease protected fragments consistently co-migrated with the probes. This could be due to either very low levels of spliced mRNAs or lack of splicing in the primary transcript of E-3. One way to circumvent the low abundance of spliced mRNAs would be to construct cDNA libraries and then determine the structure of the resulting clones.

Acknowledgements

This research was supported by the Natural Sciences and Engineering Research Council of

Canada, and by the Ontario Ministry of Agriculture, Food and Rural Affairs. We thank Paul Huber for assistance with photography.

References

Baker, C.C. and Ziff, E.B. (1981) Promoters and heterogeneous 5' termini of the messenger RNAs of adenovirus serotype 2. *J. Mol. Biol.* 149, 189–221.

Beard, C.W., Oberhauser Ball, A., Wooley, E.H. and Spindler, K.R. (1990) Transcription mapping of mouse adenovirus type 1 early region 3. *Virology* 175, 81–90.

Berk, A.J. and Sharp, P.A. (1977) Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease digested hybrids. *Cell* 12, 721–732.

Chomczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium-thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* 162, 156–159.

Cladaras, C. and Wold, W.S.M. (1985) DNA sequence of the early E-3 transcription unit of adenovirus 5. *Virology* 140, 28–43.

Cladaras, C., Bhat, B. and Wold, W.S.M. (1985) Mapping of the 5' ends, 3' ends, and splice sites of mRNAs from the early E3 transcription unit of adenovirus 5. *Virology* 140, 44–54.

Clarke, M.C., Sharpe, H.B.A. and Derbyshire, J.B. (1967) Some characteristics of three porcine adenoviruses. *Arch. Ges. Virusforsch.* 21, 91–97.

Derbyshire, J.B., Clarke, M.C. and Collins, A.P. (1975) Serological and pathogenicity studies with some unclassified porcine adenoviruses. *J. Comp. Pathol.* 85, 437–443.

Devaux, C., Caillet-Boudin, M.-L., Jacrot, B. and Boulanger, P. (1987) Crystallization, enzymatic cleavage, and the polarity of the adenovirus type 2 fiber. *Virology* 161, 121–128.

Dragulev, B.P., Sira, S., Abouhaidar, M.G. and Campbell, J.B. (1991) Sequence analysis of putative E3 and fiber genomic regions of two strains of canine adenovirus type 1. *Virology* 183, 298–305.

Garwes, D.J. and Xuan, H. (1989) Genome typing of three serotypes of porcine adenoviruses. *Intervirology* 30, 234–236.

Hirt, B. (1967) Selective extraction of polyoma DNA from infected mouse cell cultures. *J. Mol. Biol.* 26, 365–369.

Hirahara, T., Yasuhara, H., Matsui, O., Yamanaka, M., Tanaka, M., Fukuyama, S. et al. (1990) Isolation of porcine adenovirus from the respiratory tract of pigs in Japan. *Jpn. J. Vet. Sci.* 52, 407–409.

Kleiboecker, S.B., Seal, B.S. and Mengeling, W.L. (1993) Genomic cloning and restriction site mapping of a porcine adenovirus isolate: demonstration of genetic stability in porcine adenovirus. *Arch. Virol.* 133, 357–368.

Kleiboecker, S.B. (1994) Sequence analysis of putative E3, pVIII and fiber genomic regions of a porcine adenovirus. *Virus Res.* 31, 17–25.

Lee, D.C., Roeder, R.G. and Wold, W.S.M. (1982) DNA sequences affecting specific initiation of transcription *in vitro* from the EIII promoter of adenovirus 2. *Proc. Natl. Acad. Sci. USA* 79, 41–45.

Linne, T. (1992) Differences in the E3 regions of canine adenovirus type 1 and type 2. *Virus Res.* 23, 119–133.

Mittal, S.K., Prevec, L., Babuik, L.A. and Graham, F.L. (1992) Sequence analysis of bovine adenovirus type 3 early region and fiber protein genes. *J. Gen. Virol.* 73, 3295–3300.

Raviprakash, K.S., Grunhaus, A., El Kholy, M.A. and Horwitz, M.S. (1989) The mouse adenovirus type 1 contains an unusual E3 region. *J. Virol.* 63, 5455–5458.

Reddy, P.S., Nagy, É. and Derbyshire, J.B. (1993) Restriction endonuclease analysis and molecular cloning of porcine adenovirus type 3. *Intervirology* 36, 161–168.

Reddy, P.S., Nagy, É. and Derbyshire, J.B. (1995a) Sequence analysis of putative pVIII, E3 and fibre regions of porcine adenovirus type 3. *Virus Res.* 36, 97–106.

Reddy, P.S., Tuboly, T., Nagy, É. and Derbyshire, J.B. (1995b) Molecular cloning and physical mapping of porcine adenovirus types 1 and 2. *Arch. Virol.* 140, 195–200.

Reddy, P.S., Tuboly, T., Dennis, J.D., Derbyshire, J.B. and Nagy, É. (1995c) Comparison of the inverted terminal repetition sequences from five porcine adenovirus serotypes. *Virology* 212, 237–239.

Roberts, R.J., Akusjarvi, G., Alestrom, P., Gelinas, R.E., Gingeras, T.R., Sciaky, D. and Pettersson, U.A. (1986) A consensus sequence for the adenovirus-2 genome. In: W. Doerfler (Ed.), *Developments in Molecular Virology*, Vol. 8, Adenovirus DNA. Martinus Nijhoff, Boston, pp. 1–51.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Shaw, A.R. and Ziff, E.B. (1980) Transcripts from the adenovirus-2 major late promoter yield a single early family of 3' coterminous mRNAs and five late families. *Cell* 22, 905–916.

Signas, C., Akusjarvi, G. and Pettersson, U. (1986) Region E3 of human adenoviruses; differences between the oncogenic adenovirus-3 and the non-oncogenic adenovirus-2. *Gene* 50, 173–184.

Sussenbach, J.S. (1984) The structure of the genome. In: H.S. Ginsberg (Ed.), *The Adenoviruses*. Plenum Press, New York, pp. 35–172.

Tuboly, T., Reddy, P.S., Nagy, É. and Derbyshire, J.B. (1995) Restriction endonuclease analysis and physical mapping of the genome of porcine adenovirus type 5. *Virus Res.* 37, 49–54.

Vrati, S., Boyle, D., Kocherhans, R. and Both, G.W. (1995) Sequence of ovine adenovirus homologs for 100K hexon assembly, 33K, pVIII, and fiber genes: early region E3 is not in the expected location. *Virology* 209, 400–408.